

# The proposed Pan-Indic script called Shiva

Abhijit Bhattacharjee  
Founder, Luna Ergonomics Pvt Ltd

## **The Unity and Diversity of Indian languages**

If you wrote the word "Bharatiya" in several Indian language – you would have essentially written the same thing – exactly the same characters from the same one alphabet. They would have only varied in how the characters looked. This would be a pleasant revelation to most Indians. That little variation has been the bane of our being unintelligible to each other and the little alienation that exists between the provinces of India. How much better would it make you feel if you were aware that all the curvaceous characters on the hoarding and road signs that you could have spotted in your trip to Hyderabad, Chennai, Bangalore, Trivandrum, Bhubaneshwar, Kolkata or wherever you go were all words that you could have easily read if you knew the character equivalents because they are always written in your own language. And this would be true not only in India, but also in the neighbouring countries. In the neons of Bangkok, film posters in Sri Lanka, in monasteries in Tibet or road signs in Myanmar and yet others. That's wonderful isn't it! All of these languages use scripts systems that have descended from a script called the Brahmi script that was reigning in the Indian subcontinent since at least the 3<sup>rd</sup> century BC and have grown different through regions for both practical and historical reasons.

India may be a land of numerous languages (how numerous really depends on who you ask) but thankfully there are only 9 scripts. Devanagari, Bengali, Telugu, Tamil, Gujarati, Kannada, Malayalam, Oriya, Gurmukhi. There is an interesting anecdote about how the scripts of India vary. It is said that in the ancient times, the North Indians wrote on Bhurjapatras and the South Indians on palm leaves. If you wrote with straight lines and sharp corners on a palm leaf with a sharp stylus, it would injure the leaf and leave it open to tearing and quicker degradation and hence the South Indian scripts are rounded and avoids sharp corners. The gradual change of this form in the scripts of India is easy to see as one gradually moves from the North to the South - from Hindi to Bengali and then via Oriya to Telugu, Kannada and finally to Tamil and Malayalam. And if you wish to explore more – into Sinhalese. This is the story behind scripts of South India looking like jalebis to North Indians. This story is also about another practical need of the times that needs us to reform our scripts for this digital century.

While we were working towards supporting Indian languages on the mobile phone we were exposed to the scripts of India in their deep details. We studied the characters and how much of it were really part of the living languages of today. And we found this unity amongst the Indian languages with small differences. Somewhere a few extra characters were used, somewhere a few less, but all part of the same unifying superstructure of Indian language alphabets – Ka, Kha, Ga, Gha..

We were also exposed to the complexity of form of the Indian language scripts and the resultant difficulties in supporting them in digital devices like the mobile phone. The languages of the Indian subcontinent and around belong to a typographically different class of scripts called the Abugida scripts for their very different rules on how words are written from characters. With its system of consonants (Vyanjan varna), independent vowels (Swar Varna), dependent vowels (matras), special symbols (chandrabinu etc) and how the matras are rendered, variously before, after, over, under or around the consonants requires complex composition and rendering routines for the devices. Then there are the unique ligatures of yuktakshars (consonant clusters) which can be very large in number for each language. Because all of this is radically different from how the other languages of the world are written, the complexity that is required to handle this correctly is poorly supported in the devices of the world. Hence there is a challenge there in supporting this well in devices. Some do, others don't and yet others do poorly.

<p style="text-align: center;">Hindi</p>	<p style="text-align: center;">Bengali</p>	<p style="text-align: center;">Telugu</p>
<p style="text-align: center;">Marathi</p>	<p style="text-align: center;">Tamil</p>	<p style="text-align: center;">Gujarati</p>
<p style="text-align: center;">Kannada</p>	<p style="text-align: center;">Malayalam</p>	<p style="text-align: center;">Oriya</p>
<p style="text-align: center;">Punjabi (Gurmukhi)</p>	<p style="text-align: center;">Assamese</p>	<p style="text-align: center;">Shiva</p>

The unity and diversity of Indian scripts. They are orthographically same, only typographically different. We can easily unite them into one script designed from the drawing board with all modern scientific considerations. This is the inspiration behind Shiva.

## Literacy

Alongside, we were also brought to ponder on whether this complexity of form has ever posed a barrier to literacy in India. The insight was brought about by my observation of a 3 year old daughter of a friend. She was put in a play school where on her way to learning the English alphabet, she was at first being drilled on drawing standing lines, sleeping lines, slanting lines. The little child was growing control on her fingers in drawing these simple shapes and very soon she was drawing these well. After three weeks of learning when I enquired with her mother, we learnt that she had mastered the whole English alphabet but was having difficulty with G. What set aside this particular character was the complexity of its shape. I wondered at what age could an Indian child be expected to learn his alphabet. Is it the age of 5. A late start in life by design?



*3 year old Pari learning how to write standing lines, sleeping lines while at nursery school stage. She was first taught L and T to make it simpler. In a few weeks she had learnt the whole English alphabet but was finding difficulty in learning to write G. If G is difficult, imagine what our Indian scripts look like.*

*At what age does our Indian kid get to learn his alphabet?*

*A late start in life by design?*



*9 year old Yash, comes first in his class but occasionally makes mistakes in writing his Hindi.*

What is true for a child is also true for a grown up. Take the case of adult education. Say, when a grown up were to learn writing alphabets or his name for the first time as an adult. If he is going to make very little progress in learning, that person is likely to give up quickly and will not turn up on the third day, whereas if learning was easy and he was making rapid progress, that would be reinforcing to his motivation to learn and stay by to learn far more and we could have made that transforming change of his life from illiteracy to literacy. India has the largest number of illiterates in the world. What stops us from reforming our languages towards simpler scripts. What practical purpose do our complex scripts serve?

We then looked at the historical context of our scripts and came across the following alarming conclusion. A much older Brahmi script was extant in our country from at least the 3<sup>rd</sup> century BC. During all the golden ages of India, the times of Panini and Pingala, Aryabhatta and Bhaskara, the Mauryas and Ashoka, India was writing in the Brahmi script. Prakrit, Pali, Sanskrit and others were being written using the much simpler Brahmi script. The Brahmi script predates the Sanskrit language in its written form. It is only as recently as 13th century that we started writing in the current Nagari script. And that period curiously matches with the period when India was enslaved by invading foreign armies and we have been slaves ever since. Did this change have a role to play in



characters were largely replaced by simplified forms, reduced number of keystrokes to aid printing, learning and this did improve literacy. Japanese went through another reform to support phonetic representation of words, sometimes western words. And the Koreans had reinvented their script from scratch on systematic lines in 1443. It was ridiculed at first by scholars but eventually it is the only script that both North and South Korea writes in – now called Hangul.

### **The proposed Pan-Indic script called Shiva**

The aim of Shiva is to support a single generic vector based script common for all Indian languages so that text can be displayed in this manner on the simplest of devices. A common script for easy interoperation between Indian languages and bring on a possibility of common display systems which would be efficient.

Shiva Script has been developed with the following design goals.

1. A common alphabet for all Indian languages which is inclusive of all of them in a consistent logical system.
2. To keep all the phonemic and orthographical rules intact of the original Abugida representation system of matras, yuktakshars etc. And to treat them in familiar ways only changing the physical form of the individual characters (the typography).
3. To develop a script which is suitable for segment display systems by using only straight lines. (16 segment displays : Form Lotus) Segment displays are low cost displays like calculators.



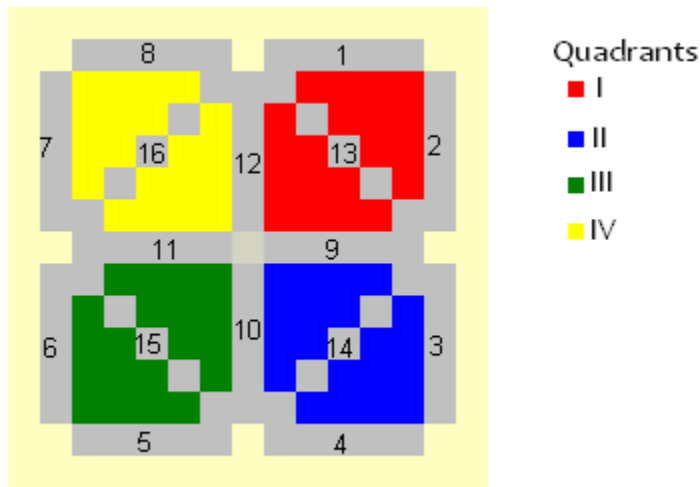
*A segment display system as above is a lightweight, low cost, low complexity display technology suitable for English but not for Indian languages who have to opt for costly alternatives like pixel based or neons, even for simplest implementations, apart from being non standard and non compatible with the worlds systems. This shall keep adding to our incremental costs and disadvantage along the century if we do not reform our script. The Shiva script will allow one railway announcement to convey in each Indian language – the delay of a train for example.*

4. To develop a script that has good recognisability of each characters and readability from far.
5. To develop a system of script which is deducible making it easy for anyone to learn and interpret. The glyphs are logical and can be arrived at through a few simple rules.



## Rules

The fundamental typographical construct for Shiva is the following 16 segment block.



The segments are numbered as shown and the block could be resolved into four quadrants.

### *Consonant forming rules*

Consonant are always part of the quadrant IV, III and I but never on Quadrant II. Quadrant II is reserved for dependent vowels (Matras) which could occur underneath consonants.

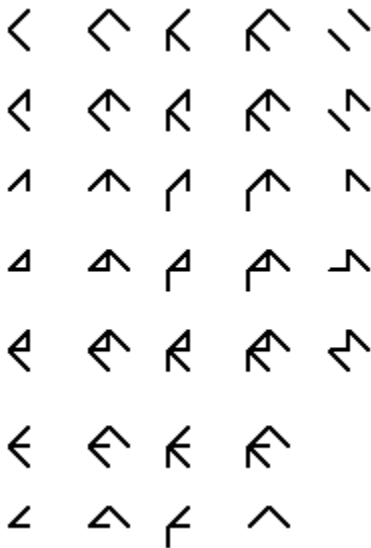
The first column of the consonants are the following basic forms.

क	<	Kantha (Guttural)
च	<	Talavya (Palatal)
ट	^	Murdhanya (Retroflex)
त	Δ	Dantya (Dental)
प	←	Osthya (Labial)
य	←	6th Row
श	←	7th Row

For the rows (below) one can see that certain additional segments have been added as indicated in red. And for the last element of the row, the absence of the common segment 16 indicated in gray.



Each row follows the same rules of construction, so a child could quickly draw all the columns and rows by just knowing the above basic column and basic row.



The segment 5 is reserved for the nukta (bindi below character) when a consonant uses it.

The Halant is indicated by the segment 1 which is reserved exclusively for it. When a consonant accompanies a halant – it indicates that the next consonant is part of the consonant cluster – a resulting yuktakshar.

### *Vowel forming rules*

The whole of Quadrant II is reversed for dependent vowels, (matras) that would be entirely contained within it. Matras form part of the consonant and would always be accommodated unambiguously within the same block.

The same dependent vowel form in the quadrant II with the additional segment 11 forms the respective independent form for that vowel (swar varna).

The vowels have two forms – chhoti, bari. The bari is always the symbol of choti plus the additional segment 14.

*Numbers are always contained within Quadrant I only.*

Special symbols as shown in Figure 1.

Extra characters used by South Indian languages follow the same rules, shown on right side of Fig 1.

The purna viram, as to indicate an end of sentence is a straight line (segment 12 and 10)

All characters and situations are covered with a unique unambiguous mapping and character types are orthogonal to each other. Some example sentences are indicated at the bottom of Fig 1.

The Shiva script is not merely wishful thinking, an aspiration or a theory. It is a fully developed technology, already inclusive in its support for all Indian languages without exceptions and is working on a million phones where it is being experienced by thousands everyday. Its efficacy is readily demonstrable to the nation.

### **Unity of India**

When Hindi was imposed as a national language, there were protests in certain parts of the country. The rationale of the protest was imposition of one on another, a learning burden on another, a perceived cultural hegemony of one. However Shiva is not one on another, it is the good fresh one for all, to the advantage of all and no group is disadvantaged. Shiva is however not about national language, it is only about national agreement on a common script. If India adopts, the subcontinent will follow. Shiva can unite us. We shall know the true extent of how we are deeply united and it will be difficult to divide us.

Such a proposition may seem absurd today, but it is again similarly absurd to imagine that India will continue well into this century without simplification reforms of this sort. It will increasingly become an inescapable need and hence it is better that we start talking about it today.

The strategy we have adopted is through inclusion of full support of the Shiva script in our Panini Keypad system which supports convenient writing and interoperation between all Indian languages, including Shiva on the mobile phone. This is a mass consumption technology meant to run in every hand. This we believe will result in curiosity and familiarity of the common man to this concept so that on a later day when such an idea would need to be imposed, we shall find informed opinion on it from the masses. It shall not appear like an alien imposition to be resisted without full consideration of its merits. Any debate on this would then be a highly informed debate.

Shiva can be included in Unicode support. The IPV6 is being defined at this time, which shall allow inclusion of Indian language words in Internet domain names. Working with Shiva will mean unity and interoperability between Indian languages in the domain name space. You would want to type in the Govt of Andhra Pradesh website even though you didn't know Telugu.

### **The need to preserve our linguistic diversity**

All of this may appear to some as an attempt to remove the linguistic diversity of India. This is definitely not the case. The wealth that we have in our languages is well recognized. Human culture is held in its languages and each are a receptacle for its own offerings and we need to preserve it. The introduction of Shiva will only mean that we shall also be able to access each others offerings more freely. Today, a Bengali poet has little opportunity to enrich himself from the works of the Tamil greats. Today the only way we interact and learn from each other is via the link language of English where much is lost in translation. Shiva will only facilitate this free interchange and vice versa. Think of the synergies that will result.

There could be some who could propose working in the roman script because a huge part of the world already works in the roman script. Before one could even consider this we wish to bring out the following about Indian language alphabets.

Indian alphabets are the most scientifically derived from very sound principles. They are classified according to the phoneme groups – guttural, palatal, retroflex, dental and labial. They are so qualitatively superior to the latin or any other alphabets of the world that they are not even worthy of a comparison. If you look at the International phonetic alphabet – a global standard arrived at in modern times, you will find that it is merely a reproduction of the work that was done in India on the subject 2500 years ago. There is also a matter called efficiency of an orthography – this is when there is one grapheme to a phoneme and vice versa. And that is the case here. The Abugida system similarly offers efficiency in typography in representational space. If there is any alphabet system which is worthy of offering leadership to the world in the future of global linguistics, it is this. While there is no need to be jingoistic about it today we must preserve and grow this for that day in the future. So it would be incorrect to abandon what we have for a less efficient historical reality.